

Submission in Response to NSF CI 2030 Request for Information

DATE AND TIME: 2017-04-06 12:24:27

PAGE 1

REFERENCE NO: 312

This contribution was submitted to the National Science Foundation as part of the NSF CI 2030 planning activity through an NSF Request for Information, https://www.nsf.gov/publications/pub_summ.jsp?ods_key=nsf17031. Consideration of this contribution in NSF's planning process and any NSF-provided public accessibility of this document does not constitute approval of the content by NSF or the US Government. The opinions and views expressed herein are those of the author(s) and do not necessarily reflect those of the NSF or the US Government. The content of this submission is protected by the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (<https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode>).

Author Names & Affiliations

- Debra Laefer - New York University

Contact Email Address (for NSF use only)

(Hidden)

Research Domain, discipline, and sub-discipline

Civil Engineering, Remote Sensing/3D Modeling

Title of Submission

Perceived Needs for Big Remote Sensing Data in Urban Contexts

Abstract (maximum ~200 words).

This response very briefly summarizes anticipated and perceived needs to advance remote sensing research and applications in urban contexts as the affiliated datasets grow ever larger and denser.

Question 1 Research Challenge(s) (maximum ~1200 words): Describe current or emerging science or engineering research challenge(s), providing context in terms of recent research activities and standing questions in the field.

The massive nature of current and emerging 3D point cloud datasets makes them increasingly challenging to manage and to query. Existing systems struggle to manage this data in a truly 3D manner and current storage, indexing, and querying techniques often lead to challenges for productive analyses. Difficulties arise in, for example nearest neighbor searches, because both direct data and contextual data related to a single object (e.g. building, road, neighborhood) are indiscriminately held in multiple locations. Furthermore, the multi-billion-point nature of many of these datasets cannot be effectively processed on standard, stand-alone machines because of their multi-terabyte size.

Question 2 Cyberinfrastructure Needed to Address the Research Challenge(s) (maximum ~1200 words): Describe any limitations or absence of existing cyberinfrastructure, and/or specific technical advancements in cyberinfrastructure (e.g. advanced computing, data infrastructure, software infrastructure, applications, networking, cybersecurity), that must be addressed to accomplish the identified research challenge(s).

Submission in Response to NSF CI 2030 Request for Information

DATE AND TIME: 2017-04-06 12:24:27

PAGE 2

REFERENCE NO: 312

To resolve the challenges outlines in Question 1, the following technical advancements will be required:

1. Scalable means to query remote sensing data sets
2. Robust means for change detection in remote sensing data sets
3. Flexible and scalable data indexing and storage concepts that would full support the plethora of existing public sector data sets including full waveform laser scanning data sets, other signal type data sets, and hyperspectral data. This must robustly address disparities in granularity and cadence and better exploit distributed computing.

Question 3 Other considerations (maximum ~1200 words, optional): Any other relevant aspects, such as organization, process, learning and workforce development, access, and sustainability, that need to be addressed; or any other issues that NSF should consider.

Despite, NEES/NEHRI data archiving requirements, to date almost nothing has been derived from this enormous data resource. This is not only wasteful but criminal, as extensive datasets remain archived but unused, requiring resources for their maintenance but providing little impact beyond the original project which drove their collection. Encouraging applicants to propose projects leveraging these existing resources would not only reduce duplication but would also make the most of the resources that have been so laboriously collected and archived. The research community would benefit greatly by more funding opportunities that specifically target the reuse data from previous NSF projects.

Consent Statement

- "I hereby agree to give the National Science Foundation (NSF) the right to use this information for the purposes stated above and to display it on a publically available website, consistent with the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (<https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode>)."
-